APPENDIX A

DEPENDENCE OF RECREATION PARTICIPATION ON BACKGROUND CHARACTERISTICS OF SAMPLE PERSONS IN THE SEPTEMBER 1960 NATIONAL RECREATION SURVEY

BY CHARLES PROCTOR

TRANSFORMATIONS AND MARGINAL TABULATIONS

Table 1 shows the numbers of cases by region, sex, and color divisions. Analysis was conducted on 3,647 unduplicated cases.

Many of the independent variables were, at this stage, in qualitative form in accordance with the code key. Table 2 indicates how the independent variables X_1 through X_{30} were constructed from the data supplied by the Census Bureau. The activity variables Y_1 through Y_{15} were constructed in each case by taking the square root of the number of different days of participation.

This square root transformation was an attempt to reduce the skewness of the original distribution so that the usual statistical tests which assume normal distributions might not be too misleading. Also it reflects my judgement that the difference between two people, one reporting no days, the other with 1 day in some activity, is not "equal" to the difference between one reporting having spent 20 days and another 21. It seems preferable to equate differences of "from none to 1" with such gaps as: "from 9 to 16" or "from 16 to 25," etc.

Table 3a shows the means and table 3b the standard deviations of the 15 activity variables for region by sex by color subpopulations.

Table 1. Numbers of cases (sample persons) behind computations for factor analysis and regression analysis and adjusted distribution of sample persons (adjusted sample persons) behind tabulations done by Census Bureau

Region and color	Sample persons			Adjusted sample persons		
	Sex		Total	Sex		Total
	Male	Female		Male	Female	
Northeast:						
White	427	528	955	536	622	1,158
Nonwhite....	28	36	64	36	45	81
Total	455	564	1,019	572	667	1,239
North Central:						
White	500	526	1,026	579	602	1,181
Nonwhite....	33	36	69	44	45	89
Total	533	562	1,095	623	647	1,270
South:						
White	422	417	839	525	507	1,032
Nonwhite....	82	120	202	108	152	260
Total	504	537	1,041	633	659	1,292
West:						
White	218	250	468	268	309	577
Nonwhite....	11	13	24	14	17	31
Total	229	263	492	282	326	608
U.S. Total	1,721	1,926	3,647	2,110	2,299	4,409

Note: The number of male and female adjusted sample persons are estimated from totals, and may differ slightly from the actual.

FACTOR ANALYSIS OF ACTIVITY INTERCORRELATIONS

If the patterns of recreation participation of two people are compared, one will find differences and, depending on his point of view, similarities. The purpose of the factor analysis is to fix on a particular point of view for determining similarities. Thus, if one of the persons camps and the other goes on nature walks they are similar from the point of view which contrasts backwoods against the more developed setting, which is one of the four factors to be introduced shortly.

The factors themselves are defined as linear combinations of the 15 participation scores. This is an operational definition assigning to each sample person a score on all four factors. The theoretical definition is turned around to read that the participation scores manifest the factor scores. Thus, knowing a person's factor scores, we can guess what activities he favors and which he tends to avoid.

The computational exercise which is used to convert the 15 by 15 correlation matrix (see table 3c) of participation scores into a matrix of factor loadings will by necessity yield "some" set of loadings.1/ But whether these numerical factor loadings identify an important, useful, or fruitful collection of factors is not guaranteed. In the present case three kinds

1/Each factor analysis began by inputting the correlation matrix "with ones on the diagonal". The machine than printed out the eigen-values and vectors after which the first four eigen-vectors [or principal axis factor loadings as they are known in factor analysis literature] became the input to a varimax rotation program. This scheme of rotation has the machine trying various directions of rotations so as to produce either very large (in absolute value) or very small (hopefully near zero) factor loadings.

Table 2. Construction of independent variables in recreation regression computations

Variable	Possible values	Definition
X_1 –Linear age	12 to 99	Age of sample person (S.P.).
X_2 –Quadratic age	144 to 9,801	Square of age of S.P.
X_3 –Cubic age	1,728 to 970,299	Cube of age of S.P.
X_4 –Rurality	0,1	Rural farm (=0), others (=1).
X_5 –SMA city	0,1	Not in SMA (=0), in SMA (=1).
X_6 –Urbanization	1 to 8	Reproduces col. 15, card 1. From: Urbanized area 3 million or more (=1) to rural (=8).
X_7 –Married	0,1	S.P. is married (=0), unmarried (=1).
X_8 –Child impedance	0 to 3	No. children in family (=0). Youngest (child or sibling) is 12 or over (=1). Youngest is 5 to 11 (=2). Youngest is under 5 (=3).
X_9 –Meaningfulness of response on employment status of S.P.	0,1	S.P. at work or looking (=0), other (=1).
X_{10}–S.P.'s occupation (status, prestige)	1 to 6	Codes: 00 to 05 (=1). 06 to 09 (=2). 10 to 13 (=3). 14 to 27 (=4). 28 to 36 (=5). 37 up (=6).
X_{11}–S.P.'s occupation (middle classness)	0,1	Codes 06 to 27 (=1), others (=0).
X_{12}–Meaningfulness of response on employment status of head	0,1	See X_9.
X_{13}–Head's occupation	1 to 6	See X_{10}.
X_{14}–S.P.'s completion of high school	0,1	High school incomplete (=0), others (=1).
X_{15}–S.P.'s education	11 to 56	From never attended (11) through elementary (31-38), high school (41-44) and college (51-56).
X_{16}–Previous farm residence	0,1	Yes (=0), no (=1).
X_{17}–Response on health	0,1	Response (=0), no response or don't know (=1).
X_{18}–Health of S.P.	1 to 4	Excellent (=1) through poor (=4).
X_{19}–Physical impairments of S.P.	0 to 2	None (=0), some (=1) to limits recreation activity (=2).
X_{20}–Per capita income in S.P. family linear		Col. 61 divided by col. 37.
X_{21}–Quadratic per capita income		Square of X_{20}.
X_{22}–Cubic per capita income		Cube of X_{20}.
X_{23}–Family income	1 to 9	Col. 61.
X_{24}–Square family income		X^2_{23}.
X_{25}–Cubic family income		X^3_{23}.
X_{26}–Marital status nonresponse	0,1	No response (=1), some (=0), zero st. dev.
X_{27}–Education nonresponse	0,1	No response (=0), some (=1).
X_{28}–Income nonresponse	0,1	Some (=0), none (=1).
X_{29}–Previous farm residence nonresponse	0,1	Some (=0), none (=1).
X_{30}–Color	0,1	White (=0), nonwhite (=1).

of evidence were used to judge the worth of the resulting factor loadings. In the first place, the measures of "variance explained," the latent roots, can be examined. These roots are helpful in deciding how many factors to attempt to identify. In the second place it was possible to compare the results with hypotheses rather well formulated in advance of the numerical work. The agreement here seems to be quite close and argues in favor of the serious acceptance of the factors. The third check consisted of a comparison of the factor patterns independently arrived at for the eight white region-by-sex subpopulations. The similarities found here lend further support to the above.

Figure 1 is a plot of the order of size of each of the latent roots of the correlation matrix (lower half of table 3) against its size. The only "rapid" drop as one moves from the first to the fifteenth root is between the first and the second. If one were to take exclusively a "rate of decrease" criterion, then he would have stopped trying to rotate or to explain factors after noting that the first principal axis one is a general factor.

However, armed with sufficient imagination and bolstered by the sample size of 3,647, the drop between the fourth and fifth roots and crossing the unit variance was taken as evidence of four factors.

Nonetheless each integer number from 3 to 11 of latent vectors were subjected to varimax rotation and the patterns examined. All of the factor patterns for five or more roots yielded at least one factor which could be interpreted as only one original

Table 3a. Means of the 15 activity variables by region, color, and sex

Vari-able	Total popula-tion	Northeast				North Central				South				West			
		White		Nonwhite		White		Nonwhite		White		Nonwhite		White		Nonwhite	
		Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
1..	0.168	0.153	0.106	0.051	0.056	0.189	0.098	0.000	0.000	0.237	0.111	0.068	0.008	0.432	0.361	0.241	0.154
2..	.632	.767	.267	.413	.139	1.004	.435	.480	.678	1.070	.388	1.136	.427	.867	.383	.273	.294
3..	.313	.284	.282	.221	0.000	.373	.326	.453	.028	.337	.190	.490	.507	.226	.396	1.414	.516
4..	.116	.065	.121	.080	.039	.110	.096	.073	.039	.149	.078	.190	0.000	.208	.191	.746	.596
5..	1.619	1.764	1.750	1.257	.874	1.734	1.980	2.005	1.535	1.393	1.454	1.172	.991	1.321	1.626	1.393	.981
6..	.437	.573	.356	.122	.166	.687	.460	.242	.096	.490	.338	.093	.025	.475	.365	.091	.077
7..	1.295	1.743	1.609	.876	.573	1.239	1.144	1.076	.437	1.325	1.120	.788	.372	1.275	1.482	1.438	.840
8..	.110	.146	.036	0.000	0.000	.157	.053	0.000	0.000	.187	.116	0.000	0.000	.173	.193	0.000	0.000
9..	.849	1.314	.701	1.216	.391	1.129	.744	2.101	.716	.881	.420	.830	.518	.803	.660	2.748	1.290
10..	.105	.144	.107	.036	.039	.096	.073	.073	.039	.121	.049	.054	.026	.192	.191	.091	.133
11..	.273	.306	.360	.172	.235	.224	.332	.159	.171	.218	.245	.117	.162	.220	.341	.339	.352
12..	.954	1.106	1.158	.832	.871	.951	1.143	.930	1.045	.687	.768	.435	.561	.866	1.083	1.038	.780
13..	.999	1.252	1.484	1.533	1.282	.722	.892	1.436	1.367	.599	.824	.746	1.260	.747	1.301	.639	.451
14..	.865	.690	.913	.411	.387	.913	1.094	1.038	.950	.686	.751	.344	.503	1.167	1.144	.840	.747
15..	.584	.705	.501	.928	.219	.700	.682	.702	.233	.563	.351	.697	.479	.597	.559	1.490	.399

Table 3b. Standard deviations of the 15 activity variables by region, color, and sex

Vari-able	Total popula-tion	Northeast				North Central				South				West			
		White		Nonwhite		White		Nonwhite		White		Nonwhite		White		Nonwhite	
		Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
1..	0.643	0.645	0.541	0.267	0.333	0.691	0.484	0.000	0.000	0.746	0.481	0.362	0.091	0.973	0.928	0.798	0.555
2..	1.210	1.457	.847	.645	.424	1.439	.948	.818	1.325	1.371	.919	1.736	1.001	1.341	.899	.467	.810
3..	1.264	1.118	1.207	.993	0.000	1.514	1.160	1.752	.167	1.425	.901	1.660	1.586	.967	1.372	3.146	1.340
4..	.615	.347	.649	.423	.236	.544	.625	.297	.236	.721	.375	.981	0.000	.745	.814	1.670	2.148
5..	2.021	2.081	2.036	1.833	1.507	2.142	2.130	2.307	2.001	2.033	1.998	1.837	1.627	1.672	1.781	1.835	2.010
6..	1.010	1.277	.962	.364	.507	1.230	.956	.849	.577	.966	.878	.449	.157	1.086	.828	.302	.277
7..	1.858	2.070	2.077	1.114	1.003	1.842	1.732	1.559	.973	1.868	1.717	1.432	.993	1.811	1.935	2.031	1.349
8..	.524	.664	.273	0.000	0.000	.590	.294	0.000	0.000	.665	.595	0.000	0.000	.624	.699	0.000	0.000
9..	1.703	2.058	1.417	1.977	.918	1.969	1.436	3.089	1.315	1.918	1.064	1.716	1.511	1.618	1.462	3.492	2.714
10..	.488	.592	.457	.189	.236	.477	.388	.297	.236	.511	.357	.289	.203	.762	.620	.302	.480
11..	.805	.999	.969	.583	.586	.733	.766	.535	.506	.731	.716	.495	.541	.841	.789	.611	.930
12..	1.121	1.254	1.287	.955	.890	1.115	1.156	1.052	1.092	.971	.889	.662	.971	.975	1.156	1.103	1.158
13..	1.864	2.115	2.182	1.927	1.896	1.729	1.717	2.338	1.903	1.415	1.705	1.581	2.089	1.581	1.988	1.500	.883
14..	1.230	1.095	1.287	.845	.748	1.288	1.309	1.709	1.115	1.139	1.187	.641	.942	1.373	1.230	1.060	.947
15..	1.106	1.261	.949	1.137	.616	1.167	1.228	1.259	.677	1.161	.922	1.223	1.079	.946	.973	1.749	.813

Table 3c. Intercorrelations over all sample persons among activity variables (in lower left-hand triangle) and mean square contingency coefficients for the same variables, when only some versus no participation is distinguished (in the upper right-hand triangle)

	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	Y_{11}	Y_{12}	Y_{13}	Y_{14}	Y_{15}	Activity
Y_1		.2265	.1072	.1453	.0263	.2189	.1764	.2020	.1517	.3054	.1381	.1323	.0760	.1456	.0104	Camping.
Y_2	.2118		.0960	.1163	.0761	.2858	.1961	.1300	.1427	.1010	.1041	.1616	.0179	.1053	.1126	Fishing.
Y_3	.1028	.0920		.2267	.1013	.1408	.2665	.1203	.3174	.1974	.2029	.1424	.1981	.1348	.1755	Bicycling.
Y_4	.0917	.0639	.1454		.0658	.1413	.1883	.1492	.2178	.1345	.1212	.1220	.1007	.1126	.1648	Horseback riding.
Y_5	.0385	.0733	.0911	.0759		.1147	.1925	.0368	.1890	.0544	.1719	.2499	.2398	.3984	.2407	Driving for pleasure.
Y_6	.2252	.3707	.1229	.1082	.1244		.3004	.3673	.2323	.1121	.1491	.1831	.1046	.1989	.2156	Boating.
Y_7	.2070	.2065	.2813	.2202	.2002	.3934		.2357	.3702	.1877	.2076	.3437	.2142	.2290	.2442	Swimming.
Y_8	.2104	.1596	.0579	.0981	.0635	.4655	.2847		.1524	.0874	.0514	.1117	.0403	.0689	.1258	Water skiing.
Y_9	.1526	.1431	.3622	.1859	.1626	.2244	.4201	.1239		.1848	.2036	.2872	.2158	.2217	.3179	Games, sports.
Y_{10}	.2786	.0840	.2167	.1183	.0578	.1524	.2284	.0905	.1902		.1943	.1255	.1348	.1415	.1175	Hiking.
Y_{11}	.1274	.1185	.1662	.0942	.1658	.1573	.2286	.0827	.1956	.1885		.2062	.2718	.2956	.1668	Nature walks.
Y_{12}	.1465	.1799	.1305	.1066	.2706	.2392	.3820	.1370	.2597	.1347	.2297		.1934	.3188	.2279	Picnicking.
Y_{13}	.0555	.0087	.1561	.0321	.2240	.0692	.1628	.0212	.1664	.1319	.2482	.1379		.2878	.1880	Walking for pleasure.
Y_{14}	.1449	.0626	.0843	.0686	.3574	.1418	.1811	.0550	.1555	.1667	.2519	.2508	.2161		.2506	Sightseeing.
Y_{15}	.0777	.1546	.1625	.1154	.2657	.1763	.2687	.0914	.3260	.1024	.1271	.2278	.1398	.1806		Outdoor events.

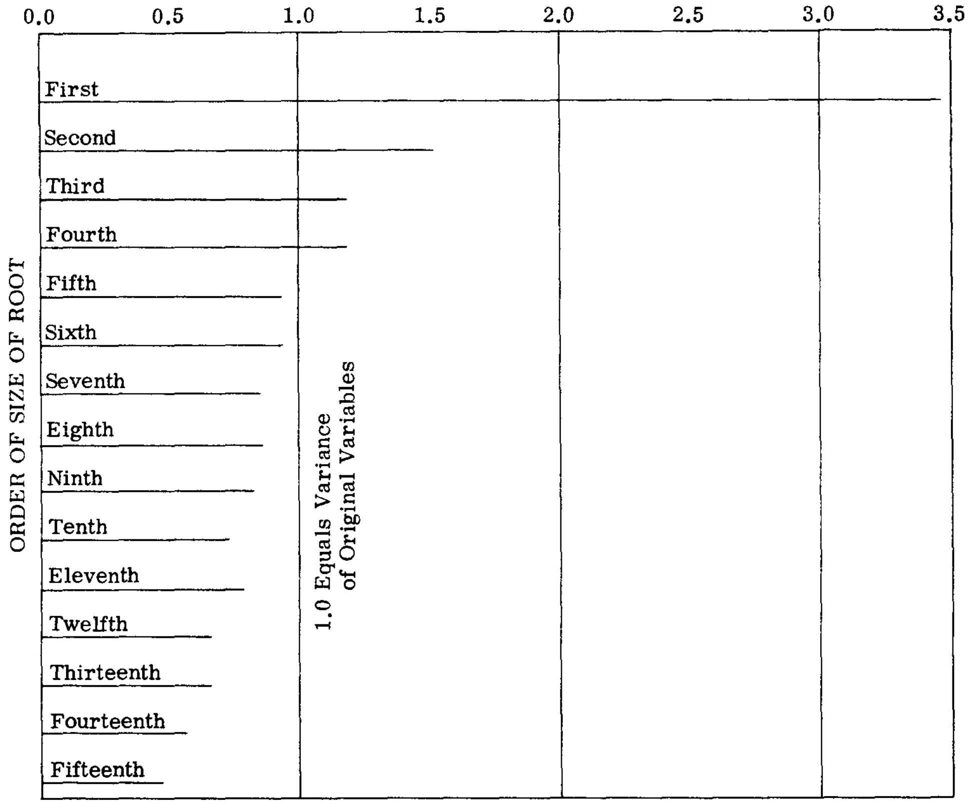


Figure 1. Size of latent roots plotted against order of size

variable. This indicated that the nonerror variance represented by the fifth, sixth, to the fifteenth vectors was specific rather than common. Consequently, the decision to use four latent vectors as the basis for varimax rotation was confirmed.[2] The principal axes (or latent vector) and varimax factor matrices for the first four factors are shown in table 4.

The first principal axes factor loadings are all positive, while there is a mixture of signs (albeit an unbalanced mixture) for all varimax factors. Notice also the close parity with respect to communality among the first three varimax factors. A single "general factor" could make sense in that people active in any one kind of outdoor recreation are people who are generally active in outdoor recreation. However, this is just not important enough in comparison with the possibility that certain people might favor one type of activity while others show a different style.

[2] A technique suggested by Bartlett and reported in M. G. Kendall, "A Course in Multivariate Analysis," Hafner, New York, 1957, pp. 95-98 for deciding if the latent roots are effectively differentiated was applied to the present data. All values to the thirteenth were found significant, so by this means we would be entitled to claim that four is modest enough. For example, $\lambda_{14}=0.50606102$ and $\lambda_{15}=0.45207721$ were found as fourteenth and fifteenth roots. The test statistic to determine if a thirteenth factor is differentiated was found to be 11.05 with a distribution under the null hypothesis of a 1 degree of freedom chi-square.

The hypotheses of styles of outdoor recreation participation were elaborated in discussions with Abbott L. Ferriss and H. Douglas Sessoms. Let

Table 4. Principal axis and varimax rotated factor loadings for total population

Activity	Factor loadings[1]							
	Principal axis				Varimax			
	1	2	3	4	P	W	A	B
Camping	41	31	-05	-51	03	39	02	61
Fishing	40	41	20	05	05	60	08	03
Bicycling	44	-11	-57	09	04	-07	69	24
Horseback riding	32	05	-34	18	-03	08	49	05
Driving for pleasure	43	-44	41	15	73	08	05	-17
Boating	58	50	25	04	12	78	14	08
Swimming	70	12	-11	17	26	43	54	09
Water skiing	41	54	22	00	-02	71	04	08
Games, sports...	61	-08	-36	28	21	13	72	04
Hiking	42	01	-35	-53	09	05	25	71
Nature walks....	47	-25	03	-32	46	05	12	40
Picnicking	57	-11	22	12	49	32	23	-03
Walking for pleasure	35	-46	03	-22	53	-16	12	26
Sightseeing.....	45	-40	35	-24	70	07	-07	20
Outdoor events..	49	-18	04	42	39	17	44	-26
Variance	3.47	1.51	1.21	1.11	2.06	2.01	1.91	1.32

[1] All entries in the body of the table are in hundredths.

me quote from notes summarizing my hunches as to styles of recreation participation:

Activity Groupings by Cultural Context

I Backwoods—The relevant standards are escape from formality of interpersonal relations and observance of rules for gaining approval which are mutually exclusive with those used in work-a-day situations. The activities which in many cases seem to answer to these role specifications are: Camping, fishing, hunting, nature walks, hiking, mountain climbing, and canoeing.

II Boat culture—Originally this was named Mobility Culture and would have covered motorcycles and sports cars, as well as boats and automobiles, but the relevant activities distinguished on the questionnaire list are simply: Boating and water skiing. Gregory Stone used the term "highway culture" to describe the features of shiny, long-tail fins, speed, showmanship, and the element of risk involved and these seem to underly the activities considered.

III Country Club to Picnic Ground Complex—These are activities of moderate to heavy public setting with rather elaborate "rules of the game" and traditions with widely distributed and agreed-upon standards of excellence. Here we include: Sailing, swimming, bicycling, horseback riding, outdoor games or sports, and picnicking.

IV Passive Pursuits—Driving for pleasure, walking for pleasure, sightseeing, attending outdoor concerts, and attending sporting events.

It is clear that the fourth activity grouping (passive pursuits) is reflected in P, the first varimax factor. The second grouping (boat culture) appears to be factor W, while the first grouping (backwoods) is varimax factor B. The third grouping and factor A may be identified. Although the correspondences are apparent, considerable reappraisal will need to be done to redefine the "cultural contexts."

Table 5 summarizes the activities which characterize each factor and the rationale by which their names were chosen is apparent from the listings.

Figure 2 is a further attempt to characterize the factors. Activities far from the center are those which clearly form part of only one factor; those nearer the center have ties to many factors.

The backwoods factor has a variance of 1.32 (see bottom table 4) which indicates that it is "roughly worth about one and one-third 'activity.'" Therefore, its status is not too well supported and, as will be seen, its reappearance among the region by sex patterns is sporadic.

The rather loosely phrased statement including "one and one-third activity" deserves comment. The correlation matrix, table 3, exhibits each activity with a variance of unity (1.0). All are standardized. The sum of the variances (i.e. 15) is conserved through the extraction of the latent vectors and roots, while the varimax rotations also preserve this total variance. This means that variability among persons with respect to the concocted variables, i.e. among their factor scores, is obtained by a reapportioning of the original variance and not by an artificial change of scale.

Personally, I have been wary of attributing much importance to factor analytic results that are new and suggestive. Too often they may be a figment of an unstable rotation problem. When an independent replication of the study again yields the same pattern naturally and without forcing, one can begin to feel confidence in the results. Consequently, I consider the following data on regional and sex comparisons of factor patterns to be of great utility, although the statistical theory used is not rigorous.

Fifteen by fifteen (activity variables) correlation matrices were computed for each of the 16 region by sex by color subpopulations. Each of the eight region by sex subpopulation matrices for the white sample persons were subjected to the factor analysis program. The matrices for the nonwhite persons were not analyzed due to the small sample sizes. The results of the eight analyses appear in table 6.[3/]

Table 7 shows the eigen-values of the five largest roots for each of the eight subpopulation correlation matrices. There is considerable uniformity among the eight sets, and it seems obvious that on the basis of variance alone the first four factors do not differ greatly from one subpopulation to the next. The Western and the Southern females show the largest sums of the first four roots indicating some greater degree of homogeneity of outdoor recreation styles there than among, say, the Northeastern or North Central females. The males are similar with respect to this sum across the four regions.

The factor patterns in each of the subpopulations were compared to the total population pattern. First, an attempt was made to locate the corresponding P, W, B, and A factors. These identifications appear in table 6. Next, a search for discrepancies between the total population loadings and the subpopulation patterns was made, and the results appear in table 8.

Among the Northeast males there was a more serious problem of correspondence of factor patterns. The first factor for Northeast males, call it 1-NEM, appears to be the W-factor with a bit more nature walks (45 loading on 1-NEM versus 05 loading on W-factor for the population) and walking for pleasure (12 versus -16) and less sightseeing (-18 versus 07).

Factor 2-NEM is a driving-picnicking factor but with more physically demanding activities represented than the P-factor has. That is, games and sports, swimming and horseback riding are prominent in 2-NEM but not in the P-factor. The other candidate for passive factor is 4-NEM, which has very low

[3/] These analyses were not contemplated in the original proposal. When the correlation matrices were first produced I asked the State College computing center at Raleigh to do the factor analyses using a quartimax program. This they generously arranged to do. Then it turned out that the data were in error due to the loss of the few cases. The results produced at Raleigh were nonetheless sufficient to show the value of such analyses and also to show the superiority of the varimax over the quartimax for this problem.

Table 5. Activities favored by persons with high scores on the factor

Factor	High	Moderate	Opposed
P-factor Passive pursuits	Driving for pleasure Sightseeing	Walking for pleasure Picnicking Nature walks Outdoor events	None
W-factor Water related activities	Boating Water skiing Fishing	Swimming Camping	Walking for pleasure
A-factor Physically demanding activities	Games, sports Bicycling	Swimming Horseback riding Outdoor events	
B-factor Backwoods activities	Hiking Camping	Nature walks	Outdoor events Driving for pleasure

total variance and on which driving is little represented (26 as compared with 73 on the P-factor).

Factor 3-NEM has high loadings on hiking and bicycling with a negative loading on driving. There are moderate loadings on camping and fishing. The importance of these contrasts among modes of transportation and the associated activities may hold up in further investigation but it is only suggestive at present. In short, the NEM factor pattern is aberrant and is worthy of further questioning but not within the confines of the present study.

In the North Central region the factor pattern for males and most of that for females conform to the population-as-a-whole pattern. For males, walking for pleasure is less a part of the P-factor than it is for females and this is true in other regions also. In the North Central region swimming is more a part of the B-factor for both sexes. The B-factor is the least satisfactorily identified of the factors, and particularly for the females. However, since its common variance is only 1.25, it is also true that, whatever the fourth factor is, it is not too important.

In the South, also, conformity to the overall pattern is noticeable especially among the males. For them, picnicking is less a part of the P-factor and more a part of the W-factor and bicycling appears to move from the A-factor to the B, relative to the population as a whole. For the females, the total population P- and B-factors cannot be too clearly identified. Fishing and picnicking occur in the B-factor of Southern females while these activities are not part of the total population B-factor. In this case the common variance, 1.96, is sizeable.

In the West both the male and the female patterns show that fishing occupies there a different position than in the population as a whole. It disappears from the W-factor and appears in the B-factor. In general the W-factor becomes more specialized—to camping and fishing.

Meaning of the correlation coefficients among activity scores

The correlation coefficients of table 3c represent a composite index of association between two activities. The mean square contingency coefficients, also shown in table 3c, reflect one component of the association. Insofar as participants (whether "light" or "heavy") in one activity are also participants in the other activity, the ϕ value or mean square contingency coefficient will be large. The correlation coefficients (r values) are large both for this reason and also insofar as, among participants, the heavy participators in one activity are also heavy participators in the other.

For example, from table 3c we find that $\phi_{34} = .2267$ while $r_{34} = .1454$. Thus, although participants in bicycle riding are somewhat more likely to be participators in horseback riding than not, these who are frequently bicyclers are not the most active horseback riders. On the other hand $\phi_{67} = .3004$ while $r_{67} = .3934$. This suggests that over and above the fact that participation itself is associated between swimming and boating the days spent in these activities are also associated.

Either of these indices, ϕ or r, could, in the sense of both computational admissibility and rationale, be used as a basis for the factor analyses. If ϕ is used, exclusive attention would be concentrated upon the act of moving from none to some participation which seems important enough from a social psychological standpoint to justify such attention. The added aspect of association between the square root of the number of days participated may be more a function of the bundles of facilities provided and other situational features of the participation than of the so-called desires of the person.

In comparision with the r values actually used in the factor analyses, the ϕ values show the following general differences:

> They are larger than r for sightseeing, walking for pleasure and horseback riding than bicycling and nature walks.
>
> They are smaller for boating, swimming, fishing, water skiing, picnicking, and driving for pleasure.

Notice the cluster of water-related activities, whose intercorrelation can be explained by the common recreation setting, with smaller ϕ values than r.

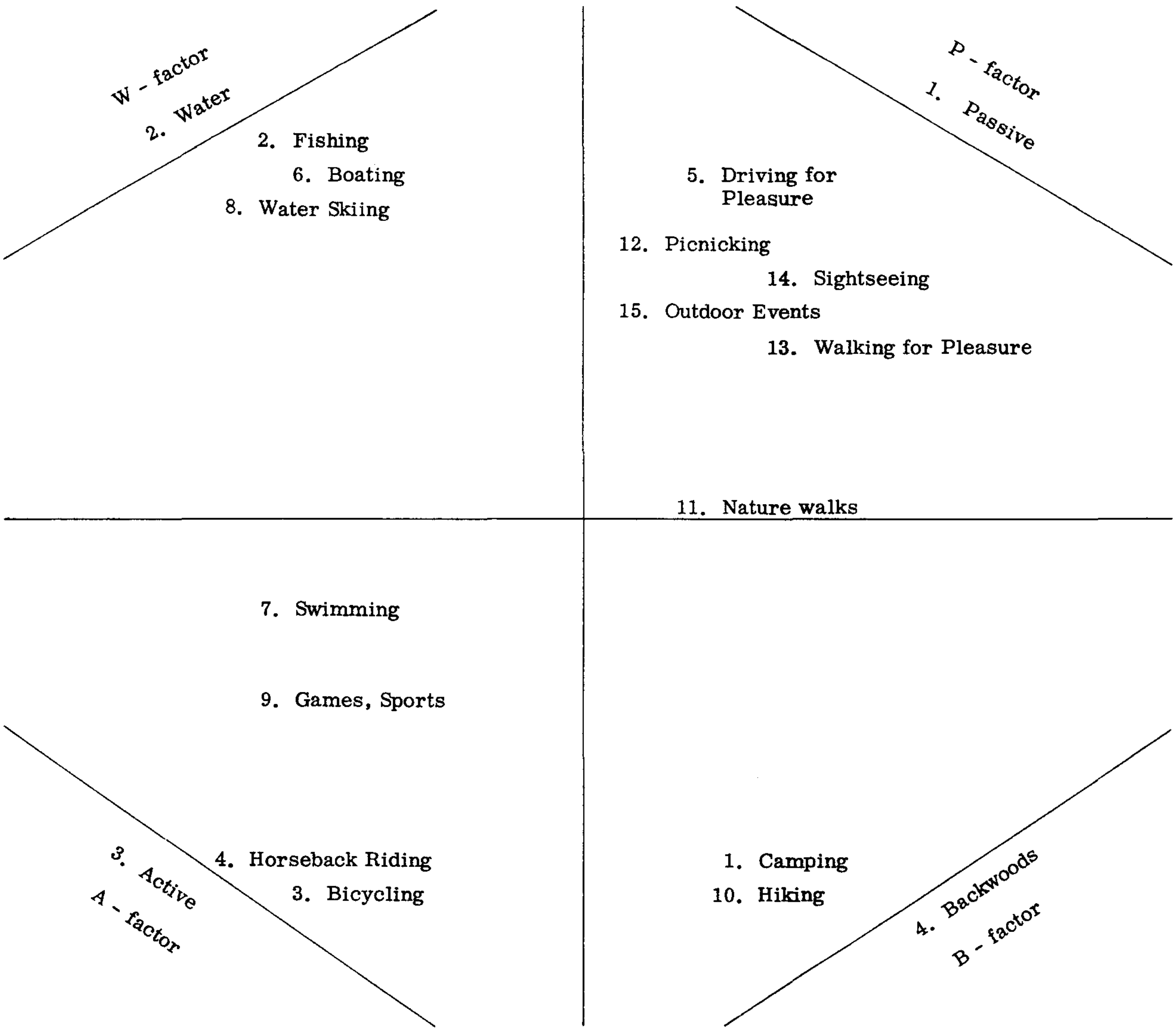


Figure 2. Pictorial representation of total population factor pattern

REGION, SEX, AND COLOR DIFFERENCES

Computation of the dependent variable scores

If X represents the 15 by N = 3,647 matrix of standardized activity scores (square roots of number of different days), then $XX'/N = R$, the 15 by 15 correlation matrix (prime denotes transpose). R appears in table 3c. Next a 15 x 15 matrix, P, was computed such that $P'R = D_{p\lambda}P'$, where $D_{p\lambda}$ is a 15 x 15 diagonal matrix of latent roots of R ordered, from top to bottom, by size. The last 11 columns of P were discarded leaving the 15 x 4 matrix, called A say, with the property $A'R = D_\lambda A'$ where D_λ is 4 x 4 diagonal and contains as elements the first four latent roots of R.

A, the so-called principal axes factor loadings, became the input to the varimax rotation routine whereby a 4 x 4 matrix T was found such that $T'T = I$ and such that the rotated factor loadings $AT = A_r$ satisfied a criterion of simple structure, namely maximization of the sum of their fourth powers. The factor loadings exhibit the correlations between the four "underlying factors" and the 15 observed variables. In order to find for each individual a score on the underlying factors we proceeded to make three trials and two errors.

At first the elements of the matrix A_r itself were employed as weights to obtain $Z' = X'A_r$ where Z is a 4 by N matrix of "underlying factor scores." However, the correlation matrix of these factor scores becomes $ZZ'/N = A_r'XX'A_r/N$, $T'A'RAT = T'D_\lambda A'AT = T'D_\lambda^2 T$. This is not in general diagonal and the underlying factor scores are thus not orthogonal.

Finally, we took as weighting matrix $B = AD_\lambda^{-1}T$. Note that these dependent variable scores, namely

Table 6. Varimax rotated factor loadings for 8 subpopulations with population factors identified (entries are in hundredths)

NORTHEAST–WHITE

Variable	Males 1	Males 2	Males 3[1]	Males 4	Females 1	Females 2	Females 3[1]	Females 4
1	21	16	44	-26	11	31	43	-15
2	65	11	30	-06	72	09	-18	11
3	11	16	67	04	-09	10	09	74
4	-12	48	27	19	-07	-10	58	26
5	-06	66	-34	26	-02	62	-12	23
6	79	15	13	00	75	06	23	08
7	54	56	17	04	46	23	30	36
8	75	06	-07	19	32	-15	55	-03
9	31	62	17	-09	22	12	31	69
10	06	02	75	20	-14	32	60	10
11	45	05	08	52	16	58	15	01
12	20	64	13	-03	30	39	-04	36
13	12	02	03	73	-04	62	08	13
14	-18	43	08	57	14	67	03	02
15	18	51	12	05	21	08	-09	62
Population factor		(See	text)		W	P	B	A
Variance	2.39	2.31	1.61	1.38	1.68	2.02	1.51	1.86

NORTH CENTRAL–WHITE

Variable	Males 1	Males 2	Males 3	Males 4	Females 1	Females 2[1]	Females 3[1]	Females 4
1	39	61	-08	08	37	01	46	09
2	59	08	19	03	48	21	-32	-06
3	21	20	-13	71	-01	12	-09	78
4	-10	04	13	57	21	-20	22	53
5	12	01	72	13	12	69	-09	-01
6	79	01	21	08	70	18	07	31
7	47	30	25	51	39	18	29	58
8	67	02	04	08	67	-01	03	06
9	16	16	21	73	04	39	02	55
10	07	71	-11	33	-09	11	77	08
11	-01	62	35	-01	20	40	16	41
12	23	07	58	08	37	52	19	12
13	-17	43	31	13	-13	60	-08	26
14	13	31	62	-02	01	67	34	-12
15	14	-22	58	43	18	50	-03	24
Population factor	W	B	P	A	W	P	B	A
Variance	2.02	1.76	2.02	1.98	1.75	2.27	1.25	1.96

SOUTH–WHITE

Variable	Males 1	Males 2	Males 3[1]	Males 4	Females 1	Females 2	Females 3	Females 4[1]
1	38	46	04	-09	-12	41	17	51
2	59	03	02	03	-08	41	01	47
3	-04	69	09	21	09	02	79	03
4	09	-10	-07	59	-05	11	77	10
5	01	-17	73	22	75	06	02	14
6	81	09	01	06	11	87	-03	09
7	35	20	09	60	49	50	33	-05
8	68	00	-03	15	-01	83	09	-06
9	04	37	11	70	44	08	65	20

See footnote at end of table.

Table 6. Varimax rotated factor loadings for 8 subpopulations with population factors identified (entries are in hundredths)–Con.

SOUTH–WHITE–Con.

Variable	Males 1	Males 2	Males 3[1]	Males 4	Females 1	Females 2	Females 3	Females 4[1]
10	01	77	-05	13	21	-13	09	69
11	14	52	36	-10	21	-05	05	71
12	64	05	27	20	19	28	23	37
13	-03	20	60	-02	64	-08	03	19
14	16	13	71	-03	46	12	08	49
15	06	03	44	48	47	11	39	03
Population factor	W	B	P	A	P	W	A	B
Variance	2.20	1.84	1.82	1.63	2.01	2.17	2.00	1.96

WEST–WHITE

Variable	Males 1	Males 2	Males 3	Males 4[1]	Females 1	Females 2	Females 3	Females 4[1]
1	29	12	06	74	-02	21	-05	72
2	25	00	22	75	15	00	18	79
3	10	-15	51	-48	00	16	52	-28
4	-09	02	57	21	41	36	-13	06
5	14	76	-14	-02	24	83	-09	-01
6	82	11	11	20	22	80	19	10
7	42	-05	64	01	-06	78	21	12
8	76	-07	-05	16	07	61	44	09
9	21	03	59	-06	14	21	81	12
10	69	19	20	-11	34	00	66	17
11	02	66	04	04	50	-01	46	17
12	38	32	33	10	63	15	28	10
13	26	29	03	-41	75	02	11	00
14	05	68	18	-01	79	04	09	-01
15	-10	32	51	11	56	12	07	03
Population factor	W	P	A	B	P	W	A	B
Variance	2.36	1.87	1.88	1.66	2.59	2.57	2.02	1.35

[1] Indicates that a sign reversal has been performed on the varimax output.

Table 7. Eigen-values of the largest 5 vectors by levels

Level	1st	2d	3d	4th	Sum	5th
Total population–White and nonwhite	3.47	1.51	1.21	1.11	7.30	0.92
Northeast (white)						
Males	3.67	1.64	1.27	1.11	7.69	1.04
Females	3.24	1.41	1.25	1.17	7.07	.99
North Central (white)						
Males............	3.81	1.50	1.31	1.17	7.79	.97
Females	3.56	1.46	1.15	1.06	7.23	1.00
South (white)						
Males............	3.25	1.64	1.42	1.18	7.49	1.06
Females	3.81	1.81	1.43	1.09	8.14	1.01
West (white)						
Males............	3.27	1.68	1.57	1.24	7.76	.98
Females	4.18	1.86	1.29	1.19	8.52	.98

Table 8. Subpopulation factor loadings relative to total population loadings (numbers in parentheses equal absolute value of difference between subpopulation and total population loadings)

	P-factor	W-factor	A-factor	B-factor
Northeast M	(See Text)			
F	Camp up (28) Hike up (23) O. Events down (31)	Camp down (28) W. Ski down (39)	Horse ride down (23)	Fish down (23) Horse ride up (53) Swim up (21) W. Ski up (47) Games up (27) N. Walk down (25)
North Central M	P. Walk down (22)	Bicycle up (28)		Swim up (21) N. Walk up (22)
F			N. Walk up (29)	Fish down (35) Bicycle down (33) Swim up (38) N. Walk down (24) Picnic up (23) P. Walk down (34) O. Events up (23)
South M	Picnic down (22)	Picnic up (32)	Bicycle down (48)	Bicycle up (45) Games up (33) O. Events up (23)
F	Swim up (23) Games up (23) N. Walk down (25) Picnic down (30) Sightsee down (24)		Horse up (28) Swim down (21)	Fish up (44) Bicycle down (21) Drive up (31) N. Walk up (31) Picnic up (40) Sightsee up (29) O. Events up (29)
West M	P. Walk down (24)	Fish down (35) Hike up (64) P. Walk up (42) O. Events down (27)		Fish up (72) Bicycle down (72) Hike down (82) N. Walk down (36) P. Walk down (67) Sightsee down (21) O. Events up (37)
F	Horse ride up (44) Drive down (49) Swim down (32) Hike up (25) P. Walk up (22)	Fish down (60) Bicycle up (23) Horse ride up (28) Drive up (75) Swim up (35)	Horse down (62) Swim down (33) W. Ski up (40) Hike up (41) N. Walk up (34) O. Events down (37)	Fish up (76) Bicycle down (52) Hike down (54) N. Walk down (23) P. Walk down (26) Sightsee down (21) O. Events up (29)

$Y' = X'B$, have covariance matrix $YY'/N = T'D_\lambda^{-1}A'XX'AD_\lambda^{-1}T/N = I$, the identity matrix. Also, as for correlations between Y's and X's, we find $\frac{1}{N} \cdot XY' = RAD_\lambda^{-1}T = AT = A_r$ the factor loadings, as is proper.

In point of fact, the activity scores were never standardized by origin, only by scale (and one further trial and error were expended to achieve this degree of standardization). This accounts for the fact that the grand means of the Y variables as they stand are not zeros. If M be taken to represent the 15 x N score matrix of unstandardized measurements and D_s be the diagonal matrix of standard deviations, then the matrix $X^* = D_s^{-1}M$ was actually used in computations in place of the X as described above.

The matrix B tells all there is to know about the dependent variables, the factor scores, but the explanation is rather unpalatable. The units are numbers of standard deviations (away from zero, rather than the mean in this case). The means and standard deviations of the factor scores for each region by sex subpopulation appear in table 9. The differences among these means exhibit the nature of the dependence of the factor scores on region and sex.

REGRESSION ANALYSIS

Within each region by sex subpopulation the pattern of dependence of the four activity scores upon the 29 other variables is exhibited in the multiple correlation coefficients and regression coefficients. These appear in tables 10 and 12.

Unfortunately, the regression program did not compute standardized regression coefficients and the independent variables were not standardized before performing the regression. Since the interpretations are to be made largely in terms of the standardized coefficients, it was necessary to compute these on a desk calculator. The number of them being large and time short, these computations were not checked, so some caution should be exercised in utilizing them.

The major problem in interpreting these regression coefficients is the fact that we tend to focus upon only one or two coefficients at a time. They were in fact computed in three "batches" or runs. The first run deleted X_9 through X_{15} and X_{20} through X_{25} the second run deleted only X_{20} through X_{25} and the final run allowed all variables to enter. Variable X_{26} was never part of the regression since it was a constant.

Regression coefficients computed in table 12 always refer to the regression equation for which the fewest variables possible appear. Thus although three estimates of the coefficient of X_1 in the equation for Y_3, say, are available we will only use the coefficient estimated by the first run where X_1 is one of 16 other variables. It also appeared as one of 23 and of 29 variables.

The sign and size of a regression coefficient thus indicate what peculiar contribution should be added or subtracted and how much, depending on a person's X-score so as to best predict or guess his Y-score, and this is a contribution over and above what contributions other X-scores of the same person make to the guessed Y-score. In a "linear" world the regression coefficients show how much to be added or subtracted for a given X_6 score, or separation from urban centers, say, no matter what age, occupational status or whatnot the person is. Thus, the statement "all other things equal" has its proper meaning if, in fact, the factor scores can be represented as linear functions of the X-scores. The rather low proportions of variance removed seem to suggest that this is not the case.

A further precaution in interpreting the coefficients or perhaps another way of saying the same thing is a hypothesis somewhat facetiously offered by M. G. Kendall "that if enough variables are used in a regression system every system is almost collinear, and consequently the matrix determining the coefficients is ill-conditioned." 4/ It is not clear as to how many is "enough", but surely 29 is pushing that boundary line.

Table 10 of percentages of variance removed yields the rather distressing finding that although the rate of "explanation of variance" may rise to about 40 percent for Y_3, it is usually within 10 to 20 percent. These percentages have an appealing interpretation as a ratio of number of days explained by number of days of variation in the population recreation pattern. The Y variables are in units of square roots of different days of participation and their variances are thus in days of participation.

Roughly speaking, if "30 percent of the variance is removed" this is saying that although knowing the population pattern of participation would allow us to be able to guess a person's number of days of participation to within, say, 10 days with a certain rate of correctness we could, by using the background information on him as represented by the independent variables, guess his participation to within (100-30) percent of 10, that is 7 days, with the same success. Thus, knowledge of the variables of group I (age primarily) allows one to reduce the uncertainty, expressed in the form of a confidence interval, about a North Central female's active factor score (Y_3) to $100\sqrt{1 - .2872}$ percent = 84.43 percent of its former length, or expressed as a confidence interval concerning number of days of "preparedness"

4/ In a discussion reported on page 58 of a paper by H. Wold, "Journal of the Royal Statistical Society," series A:119:28.

Table 9. **Means and standard deviations of the 4 factor scores for region-by-sex populations**

		Northeast		North Central		South		West	
		Male	Female	Male	Female	Male	Female	Male	Female
Y_1	Mean	1.04	1.16	.97	1.20	.68	.83	.89	1.09
	Standard deviation	.98	1.05	1.01	1.05	.89	.92	.92	.99
Y_2	Mean	.58	.26	.70	.38	.65	.29	.66	.45
	Standard deviation	1.21	.73	1.11	.73	1.08	.90	1.18	1.11
Y_3	Mean	.52	.33	.42	.27	.41	.18	.33	.33
	Standard deviation	.97	.93	1.15	.90	1.10	.79	1.07	1.11
Y_4	Mean	-.07	.03	-.12	-.13	-.01	-.03	.27	.31
	Standard deviation	1.06	.90	1.08	.79	1.06	.73	1.28	1.24

Table 10. Percentages of variance as represented by squares of multiple correlation coefficients by three groups of independent variables:
Group I —Everything except . . .
Group II —Occupation and education (X_9 through X_{15}) and
Group III—Income (X_{20} through X_{25})

Dependent variable	Source	Northeast		North Central		South		West	
		Male	Female	Male	Female	Male	Female	Male	Female
Y_1	Group I alone	3.26	10.56	14.83	15.49	8.72	16.91	5.22	11.31
	Group II over Group I	4.12	3.91	1.93	3.05	3.45	6.81	3.41	8.02
	Group III over I and II	1.00	2.98	1.47	1.39	1.52	4.32	2.32	1.61
	Total removed	8.38	17.45	18.23	19.93	13.68	28.04	10.95	20.94
Y_2	Group I alone	18.58	6.81	15.67	9.87	12.92	9.69	20.31	14.54
	Group II over Group I	4.87	3.59	2.94	1.12	3.75	.42	2.10	4.38
	Group III over I and II	3.04	3.97	2.10	2.16	.94	3.73	3.93	3.96
	Total removed	26.48	14.36	20.71	13.14	17.61	13.83	26.34	22.88
Y_3	Group I alone	35.28	31.27	35.07	28.72	32.02	27.83	41.01	33.28
	Group II over Group I	.84	1.11	3.57	1.19	1.98	.80	2.51	2.61
	Group III over I and II	.42	.95	.47	1.03	1.53	1.00	1.26	1.84
	Total removed	36.54	33.33	39.11	30.94	35.53	29.63	44.78	37.74
Y_4	Group I alone	8.78	2.79	5.41	3.47	9.91	1.66	9.82	9.84
	Group II over Group I	1.46	3.08	1.97	2.95	1.97	.28	4.29	3.35
	Group III over I and II	3.07	2.27	.80	2.67	.23	1.66	5.81	2.52
	Total removed	13.31	8.14	8.18	9.09	12.11	3.60	19.93	15.71

or "proneness" or whatnot for active outdoor recreation pursuits, it would be reduced to 71.28 percent of its former length.

Table 11, the results of significance tests of variances removed, gives, so to speak, a signal to go ahead and examine regression coefficients within certain groups of independent variables in certain subpopulations for certain factor scores. Table 12 shows the results of this search by reporting standardized regression coefficients whenever warranted by table 11 and by an associated t value of over 1.96.

Patterns of Dependence: Y_3 the Active Factor

Factor three, the active factor, is dependent almost exclusively on age which presumably indexes physical fitness, absence of cultural inhibitions and a constellation of tensions favorable to such activities. The only other variables of somewhat universal influence are X_8, called child impedance, an evident misnomer in this instance, and X_{16}, no previous farm residence.

In the South, particularly, previous farm residence seems to inhibit the active pursuits while, in the North Central region particularly, the presence of younger children encourages them. Among males in the North Central and the South a sample person being employed effects some restriction on his participation in activities of this type. In the West nonwhite persons had higher levels of Y_3 than whites.

Other evidences of dependence are more or less isolated, such as:

Males on rural farms in the North Central region had a higher level of Y_3 than those not on rural farms.

In the Northeast, males in SMSA's had higher levels of Y_3 than those not in SMSA's.

Among males in the North Central region those with the head of household working and those with less than a high school education had higher levels of Y_3 than other North Central males.

Among southern males poorness of health was associated with lower levels of Y_3.

Since the signs and sizes of the partial regression coefficients for Y_3 on X_1, X_2, and X_3 are quite similar from one to the other subpopulation. Figure 3 exhibits only one specimen curve, that for South females.

The scale along which Y_3 is measured is in terms of standard deviations so that 40 is a rather extreme quantity and in itself rather meaningless as is frequently the case with predictions for extreme values of the independent variable. It does suggest, however, that the distribution of $\hat{Y}_3$ values is quite positively skewed, and also cautions against the too literal interpretation of the prediction equation. Remember that the entire equation involves all other variables in group 1 and that the Y_3 scores represent only the additional adjustment due to age over or below that predicted on the basis of the other variables. The strategy is however quite clear—if the sample person is young throw caution to the winds and guess that his active factor score is high.

Dependence Pattern of Y_4 : the Backwoods Factor

The pattern of dependence of Y_3 is clearer than for others of the Y's, while that for Y_4 is muddier. The presence of the income variables is indicated although at a barely significant level. The interesting feature

Table 11. Significance of regression removal by 3 groups of independent variables

Dependent variable	Groups of independent variables	Northeast		North Central		South		West		No. of No's
		Male	Female	Male	Female	Male	Female	Male	Female	
Y_1	I. All else except–	No	**	**	**	**	**	No	*	2
	II. Occupation and education. .	**	**	No	**	**	**	No	**	2
	III. Income	No	**	No	No	No	**	No	No	6
Y_2	I. All else except–	**	**	**	**	**	**	**	**	0
	II. Occupation and education. .	**	**	**	No	*	No	No	No	4
	III. Income	**	**	*	*	No	**	No	No	3
Y_3	I. All else except–	**	**	**	**	**	**	**	**	0
	II. Occupation and education. .	No	No	**	No	*	No	No	No	6
	III. Income..................	No	No	No	No	No	No	No	No	8
Y_4	I. All else except–	**	No	*	No	**	No	No	*	4
	II. Occupation and education. .	No	**	No	*	No	No	No	No	6
	III. Income..................	*	**	No	*	No	No	*	No	4

of this dependence is that among West males, where the coefficients are quite large, the linear relationship is positive—the higher the income the higher the backwoods factor score. However, in the Northeast and among North Central females this is not the case.

The smallness of some of the linear income coefficients there suggests that they are responding to happenings in the extremes of the income distribution more so than to the whole range. This is true of the Northeast females family income coefficients, which suggest that very high family incomes lead to a decrease in levels of Y_4. This is contrasted with the Northeast males where the cubic coefficient is small and where high incomes characterize a resurgence of the backwoods factor since it is the quadratic coefficient which operates.

In instances where they are statistically significant the coefficients of Y_4 on age tell a story similar to the dependence of Y_3 on age, albeit on a flatter curve. The isolated variables with statistically significant coefficients can, of course, be interpreted but were quite unexpected to me. Child impedance once again appears to be an inappropriate title since southern males seem more likely to have high backwoods scores with younger children about. Completion of high school among North Central females seems tied to lower Y_4 scores. Among Northeast females the presence of a head of household who is other than at work is associated with higher backwoods factor scores, and for Northeast males those in SMSA and those in less urbanized settings have lower backwoods scores.

Pattern of Dependence of Y_1: the Passive Pursuits Factor

The major variable affecting Y_1 scores appears to be education—either as X_{14}, completion of high school, or X_{15}, educational status in general. The more educated persons have higher passive pursuits scores. The members of the health complex, X_{18} and X_{19}, also have significant coefficients in the not surprising direction—poorer health goes with less passive pursuit activities.

Age appears among the North Central region persons as positively related to Y_4 scores and a less urban setting tends to less passive pursuits, also in the North Central region. To complete the North Central picture a no response among females on income is associated with a lower passive pursuits score. This along with the same case in the Northeast region suggests that the Y_1 score may be to some extent a function of response conditions. Perhaps the education effect is also acting upon these response conditions and so upon Y_1, rather than directly upon Y_1.

Also among Northeast females income appeared to be significantly related to Y_1 and the negative nature of this relationship is apparent from the coefficients. The size of the cubic coefficient suggests the tortured path of a regression curve trying to be flat throughout most of its range with a twist at either extreme. Also among Northeast females those for which the head of household was other than at or looking for work and of lower X_{10} (occupation status variable) had higher Y_1 scores, while all other things equal, the nonwhites had lower levels of Y_1 scores than the whites among Northeast females.

For southern females the statistically notable coefficients suggest that:

Being not on a farm depresses the Y_1 score as does being further from centers, having younger children about, and being of a moderate occupational status while a head of household who is other than looking for or at work elevates the passive pursuits score.

Among western females moderateness of status is tied to higher Y_1 scores while the S.P. other than looking for or at work also leads to high Y_1 scores.

Pattern of Dependence of Y_2: the Water Related Factor

Investigation of the details of this pattern is left for the reader. There is a fertile field for exploration as the liberal sprinkling of asterisks in table II, over Y_2 cells, attests. The more universally apparent determinants of Y_2 scores seem to be:

Color—nonwhites have lower scores.

Urbanization—those away from urban centers have higher scores.

S.P. occupational status—high status males have higher scores.

Table 12. Standardized regression coefficients for cases of declared statistical significance

[Note: Coefficients of variables X_1 through X_8, X_{16} through X_{19} and X_{27} through X_{30} are partial on one another, coefficients of variables X_9 through X_{15} are partial on these previously removed and one another, while coefficients of X_{20} through X_{25} are partial on all other variables]

Dependence pattern of Y_1

Independent variable	Group	Northeast		North Central		South		West	
		Male	Female	Male	Female	Male	Female	Male	Female
Linear	I X_1	....	....	1.228	1.326	....	....	....	1.944
Quadratic	I X_2	....	....	-2.536	-3.115	....	....	....	-3.655
Age cubic	I X_3	....	....	1.216	1.551	....	....	....	1.709
Opposite of rurality	I X_4	....	....	....	....	-.110	-.089	....	....
SMSA' city	I X_5	....	....	....	....	....	....	....	....
Urbanization	I X_6	....	....	-.178	-.150	....	-.145	....	....
Unmarriedness	I X_7	....	....	....	....	....	....	....	....
Child impedance	I X_8	....	....	....	....	....	-.108	....	....
S.P. other than at work	II X_9	....	.262	....	....	....	....	....	.381
S.P. occupation status	II X_{10}	....	-.228	....	....	....	....	....	....
S.P. occupation moderateness of status	II X_{11}	....	....	....	....	....	-.179	....	.192
Head other than at work	II X_{12}	....	....	....	....	....	.154	....	....
Head's occupation status	II X_{13}	....	....	....	....	....	....	....	....
S.P. completed high school	II X_{14}	.216	....	.128	....	....	....	....	....
S.P. education status	II X_{15}	....	.143	....	.177	.212	.176	....	.300
No previous farm residence	I X_{16}	....	....	....	....	....	....	....	....
No response or DK on health	I X_{17}	....	....	....	....	....	....	....	....
Poorness of health	I X_{18}	....	-.184	-.196	-.100	....	-.221	....	....
Presence of impairments	I X_{19}	....	....	.119	....	....	....	....	.188
Per capita income:									
Linear	II X_{20}	....	-.009	....	....	....	....	....	....
Quadratic	III X_{21}	....	.437	....	....	....	....	....	....
Cubic	III X_{22}	....	-.408	....	....	....	....	....	....
Family income:									
Linear	III X_{23}	....	-.531	....	....	....	....	....	....
Quadratic	III X_{24}	....	1.558	....	....	....	....	....	....
Cubic	III X_{25}	....	-.985	....	....	....	....	....	....
Education no response	I X_{27}	....	....	....	....	....	....	....	....
Income no response	I X_{28}	....	-.104	....	-.096	....	....	....	....
Previous farm residence, no response	I X_{29}	....	....	....	....	....	....	....	....
Color	I X_{30}	....	-.118	....	....	....	....	....	....

Dependence pattern of Y_2

Independent variable	Group	Northeast		North Central		South		West	
		Male	Female	Male	Female	Male	Female	Male	Female
Linear	I X_1	-2.145	1.736	....	....	....	....	....	....
Quadratic	I X_2	3.328	-3.978	....	....	....	....	....	....
Age cubic	I X_3	-1.585	2.115	....	....	....	....	....	....
Opposite of rurality	I X_4	....	....	-.111	....	....	....	....	....
SMSA' city	I X_5	....	....	....	....	....	.171	.254	.185
Urbanization	I X_6	....	....	....	.160	.152	.128	....	....
Unmarriedness	I X_7	....	....	....	....	....	....	....	....
Child impedance	I X_8	....	-.105	....	....	....	....	....	....
S.P. other than at work	II X_9	....	....	-.195	....	....	....	....	....
S.P. occupation status	II X_{10}	.437	....	.198	....	.326	....	....	....
S.P. occupation moderateness of status	II X_{11}	....	....	....	....	.122	....	....	....
Head other than at work	II X_{12}	....	.157	....	....	....	....	....	....
Head's occupation status	II X_{13}	-.408	-.158	....	....	....	....	....	....
S.P. completed high school	II X_{14}	....	....	.163	....	.158	....	....	....
S.P. education status	II X_{15}	....	....	....	....	....	....	....	....
No previous farm residence	I X_{16}	....	....	....	....	....	....	....	....
No response or DK on health	I X_{17}	....	....	....	....	....	....	....	....
Poorness of health	I X_{18}	....	....	-.112	-.145	....	-.103	....	-.163
Presence of impairments	I X_{19}	....	....	....	....	....	....	....	....

Table 12. Standardized regression coefficients for cases of declared statistical significance–Con.

Dependence pattern of Y_2–Con.

Independent variable	Group	Northeast		North Central		South		West	
		Male	Female	Male	Female	Male	Female	Male	Female
Per capita income:									
Linear	III X_{20}	-.133	.207	.021	.296	....	.186	....	....
Quadratic	III X_{21}	.411	-.364	-.298	-.811	....	-.182	....	....
Cubic	III X_{22}	-.304	.155	.265	.532	....	.051	....	....
Family income:									
Linear	III X_{23}	.973	.135	.300	-.510	....	-.561	....	....
Quadratic	III X_{24}	-2.432	-.318	.199	1.493	....	1.852	....	....
Cubic	III X_{25}	1.637	.401	-.372	-.918	....	-1.194	....	....
Education no response	I X_{27}	....	....	....	....	....	....	....	....
Income no response	I X_{28}	....	....	....	....	....	....	....	....
Previous farm residence no response	I X_{29}	....	....	.123	....	....	....	.225	....
Color	I X_{30}	-.101	-.089	-.149	....	-.114	-.165	....	....

Dependence pattern of Y_3

Independent variable	Group	Northeast		North Central		South		West	
		Male	Female	Male	Female	Male	Female	Male	Female
Linear	I X_1	-3.630	-3.798	-4.180	-3.536	-3.535	-4.770	-4.417	-4.639
Quadratic	I X_2	6.066	6.581	6.853	6.008	6.085	8.560	7.330	6.045
Age cubic	I X_3	-2.893	-3.238	-3.119	-2.874	-2.894	-4.209	-3.416	-2.833
Opposite of rurality	I X_4	....	....	-.085	....	....	....	....	....
SMSA' city	I X_5	.130	....	....	....	....	....	....	....
Urbanization	I X_6	....	....	....	....	....	....	....	....
Unmarriedness	I X_7	....	....	....	....	....	....	....	....
Child impedance	I X_8	....	.142	.098	.147	....	.104	....	....
S.P. other than at work	II X_9	....	....	.335	....	.195	....	....	....
S.P. occupation status	II X_{10}	....	....	....	....	....	....	....	....
S.P. occupation moderateness of status	II X_{11}	....	....	....	....	....	....	....	....
Head other than at work	II X_{12}	....	....	-.128	....	....	....	....	....
Head's occupation status	II X_{13}	....	....	....	....	....	....	....	....
S.P. completed high school	II X_{14}	....	....	-.133	....	....	....	....	....
S.P. education status	II X_{15}	....	....	....	....	....	....	....	....
No previous farm residence	I X_{16}	....	....	.145	....	.144	.089	....	....
No response or DK on health	I X_{17}	....	....	....	....	....	....	....	....
Poorness of health	I X_{18}	....	....	....	....	-.112	....	....	....
Presence of impairments	I X_{19}	....	....	....	....	....	....	....	....
Per capita income:									
Linear	III X_{20}	....	....	....	....	....	....	....	....
Quadratic	III X_{21}	....	....	....	....	....	....	....	....
Cubic	III X_{22}	....	....	....	....	....	....	....	....
Family income:									
Linear	III X_{23}	....	....	....	....	....	....	....	....
Quadratic	III X_{24}	....	....	....	....	....	....	....	....
Cubic	III X_{25}	....	....	....	....	....	....	....	....
Education no response	I X_{27}	....	....	....	....	....	....	....	....
Income no response	I X_{28}	....	....	....	....	....	....	....	....
Previous farm residence no response	I X_{29}	....	....	....	....	....	....	....	....
Color	I X_{30}	....	....	....	....	....	....	.194	.134

Dependence pattern of Y_4

Independent variable	Group	Northeast		North Central		South		West	
		Male	Female	Male	Female	Male	Female	Male	Female
Linear	I X_1	-3.480	....	-2.769	....	-2.880	....	....	....
Quadratic	I X_2	6.801	....	5.069	....	5.453	....	....	....
Age cubic	I X_3	-3.377	....	-2.419	....	-2.719	....	....	....
Opposite of rurality	I X_4	....	....	....	....	....	....	....	....
SMSA' city	I X_5	-.151	....	....	....	....	....	....	....

Table 12. Standardized regression coefficients for cases of declared statistical significance–Con.

Dependence pattern of Y_4–Con.

Independent variable	Group	Northeast		North Central		South		West	
		Male	Female	Male	Female	Male	Female	Male	Female
Urbanization	I X_6	-.141	....	....	....	....	....	....	.189
Unmarriedness	I X_7	....	....	....	....	....	....	....	....
Child impedance	I X_8	....	....	....	....	.102	....	....	....
S.P. other than at work	II X_9	....	....	....	....	....	....	....	....
S.P. occupation status	II X_{10}	....	....	....	....	....	....	....	....
S.P. occupation moderateness of status	II X_{11}	....	....	....	....	....	....	....	....
Head other than at work	II X_{12}	....	.177	....	....	....	....	....	....
Head's occupation status	II X_{13}	....	....	....	....	....	....	....	....
S.P. completed high school	II X_{14}	....	....	....	-.202	....	....	....	....
S.P. education status	II X_{15}	....	....	....	....	....	....	....	....
No previous farm residence	I X_{16}	....	....	....	....	....	....	....	....
No response or DK on health	I X_{17}	....	....	....	....	....	....	....	....
Poorness of health	I X_{18}	....	....	....	....	....	....	....	....
Presence of impairments	I X_{19}	....	....	....	....	....	....	....	....
Per capita income:									
Linear	III X_{20}	-.086	-.278	....	-.003	....	....	.616	....
Quadratic	III X_{21}	.825	.873	....	.201	....	....	-1.700	....
Cubic	III X_{22}	-.703	-.584	....	-.222	....	....	1.061	....
Family income:									
Linear	III X_{23}	-.323	-.075	....	-.359	....	....	.689	....
Quadratic	III X_{24}	.246	.594	....	.782	....	....	-1.408	....
Cubic	III X_{25}	-.079	-.495	....	-.277	....	....	.949	....
Education no response	I X_{27}	....	....	....	....	....	....	....	....
Income no response	I X_{28}	....	....	....	....	....	....	....	....
Previous farm residence no response	I X_{29}	....	....	....	....	....	....	....	....
Color	I X_{30}	....	....	....	....	....	....	....	....

Poorness of health—those with poor health have lower scores.

Income—the picture is mixed.

Notice the peculiar pattern among the Northeast males with regard to both occupational status and income. The coefficient for X_{10} is positive and that for X_{13} is negative; thus the sample person with a relatively high occupational status for whom the head of household has a relatively low occupational status is guessed to have a high Y_2 score. On income likewise there is a reversal of direction from per capita to family. The reasons for these two oddities are probably somewhat distinct.

SUMMARY AND CONCLUSIONS

As in most statistical analyses, the results of this one should point out critical or crucial differences and similarities from person to person. Discovery of the similarities allows for more efficient summaries of the otherwise confusing details, while the recognition of differences suggests or confirms the presence of paths of causal influence. The summarizing portion of the analysis, the factor analysis, was rather extensively commented on earlier, but a few words of reminder may not be wasted before resuming work with the regression analysis.

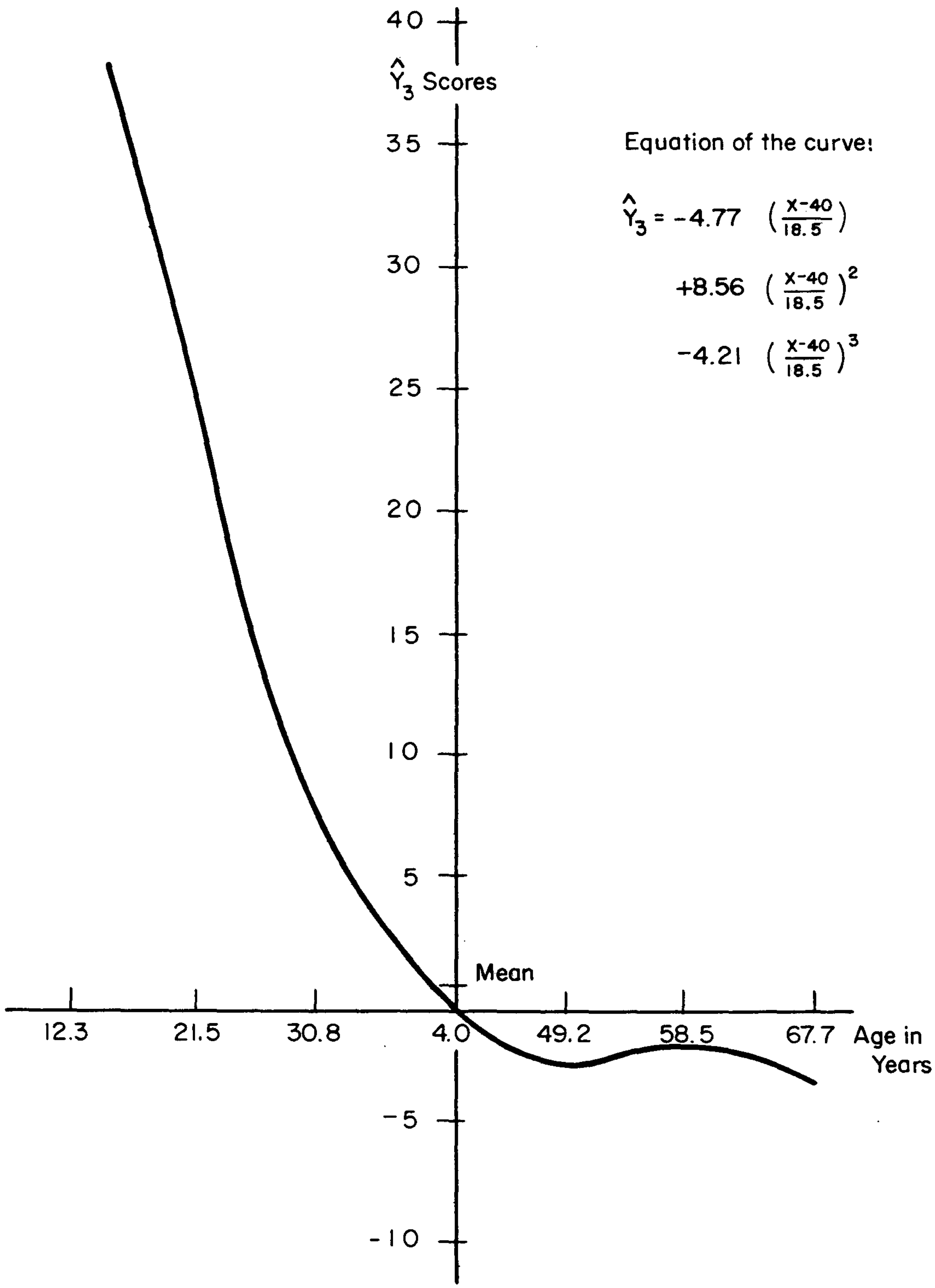
On section II. The appearance of the same four aspects of the pattern of reported recreation activity, recognizable as passive pursuits, water related, active and backwoods factor scores, was quite consistent even in the face of moving the analysis from region to region and accross the sexes. Thus, we advocate and illustrate the use of this four-part description of a person's pattern of recreation activity.

Nonetheless, it must be remembered that only 50 percent of the variance provided by the original 15 activity scores has been represented by the 4 factors and there were only 15 activities included for which the time span covered is only the 3 summer months. One should also remember that there were peculiarities in factor pattern in each region (section II) and that the correlation coefficient is not the only index of association with a legitimate claim to initiate a factor analysis.

If one person has a low score on, say, the second factor while another has a high score, we know that they differ with respect to their reports of activities during the summer. In particular the greater is the difference in Y_2 score the more likely is it that differences between the persons will be found in water related activities. It is these differences, then, which become the starting point for the regression analysis.

On section III. The sex differences were most pronounced for the Y_2 scores, the regional differences most extreme for Y_4, and the region by sex interaction was apparent in Y_3. In particular, the males showed considerably higher water related activities factor scores than the females in all regions. Although the North Central region was higher and the Northeast region lower in average Y_2 scores the regional differences were moderate.

Portion of Predicted Y_3 Scores Derived from X_1, X_2, and X_3



APPENDIX A - FIGURE 3

The backwoods factor scores were highest in the West and lowest in the North Central with the other two regions on a par closer to the North Central than the western level. The passive pursuits scores also differed quite a bit from region to region, ranging from highest in the North Central and Northeast, lower in the West and lowest in the South. Females in all regions had higher average Y_1 scores than males.

Although females have generally lower average active factor scores than males, the regional peculiarities in this respect are interesting. The difference is most pronounced in the South, nearly as much in the Northeast, less in the North Central, and for practical purposes absent in the West. The regional averages, however, show that the highest average occurs in the Northeast with the other regions near one another and arranged from North Central through the West to the South.

The extent to which the recognition of differences such as these can or should be used to forsee future patterns of use of outdoor recreation facilities is not too clear. One frequent application of such bald statements is as a check upon one's own impressions and hypotheses about recreation participation. For example, that backwoods scores are highest in the West is not surprising to me but that they are lowest in the North Central region is. Having been reared in Michigan, I know that camping opportunities are extensive in the North Central region, but now I suspect that opportunities alone are by no means a sufficient explanation of levels of participation of this kind.

The region by sex interaction for the Y_3 scores suggests that the sex role in active recreation may well differ quite a bit from region to region. There is agreement, however, from region to region that women engage in more passive pursuits and men in more water related activities. It is this last finding which I would like to check more carefully.

This suggests another use of the findings—that of serving as a guide to more careful examination of the original detailed data. Such broad patterns as have been noticed with regard to the four factor scores may be verified further and amplified by looking at table 3a or the original tabulations.

On section IV. After having applied the elaborate regression computations, it is apparent that the technique is helpful in performing a screening operation on all causal connections leading from independent to dependent variables but falls far short of being very precise concerning the amount of influence. In short, table 11 is useful but table 12 is not so much so.

The finding that Y_2, water related, and to a lesser extent Y_4, backwoods, and Y_1, passive, but not Y_3, active, are influenced by income levels confirms my own beliefs and perhaps others about the way income might influence recreation participation. But a glance at the regression coefficients for predicting or explaining Y_2, say, on X_{20} through X_{25} confuses the issue. The pattern of signs of the linear, quadratic and cubic components sometimes appears as +, -, + and at others -, +, - as one goes from region to region, from males to females and from per capita to family income.

The reason in computational language is "colinearity" as mentioned by Kendall, and quoted previously. In behavioral terms the reason is that the independent variables do not act independently of one another. They are inhibited or encouraged by concurrent states of affairs and influence one another as well as the dependent variable, if this choice of "deus ex machine" description be pardoned.

For Northeast males the dependence pattern of Y_2 on per capita income has signs -, +, - and on family income +, -, +. One is tempted to say that, since the "significance" of the coefficients for family income is greater, boating is primarily encouraged by the greater resources of the whole family while the presence of a high per capita income and what it implies about the family are actually a dampening influence on boating. This is obviously extending the guesswork beyond these data themselves but does at least raise the question.

An interesting and important feature of this particular dependence pattern are the coefficients on the sample person's occupational status. Recall that these are computed from a regression run (group II) in which income did not figure. I am quite sure that in explaining these coefficients, the fact that income and occupational status are interrelated is quite relevant.

These coefficients show that water related scores are higher when the head's occupation is high while if the sample person's occupation is lower, his Y_2 score is higher. The sample person's occupation will differ from the head's occupation only if he is not also the head and this means usually that he is younger. Of course, the age effect on Y_2 (an inverse effect) has been partialed out, so the occupational coefficients .437 (on S.P. occupation status) suggests that:

If two persons, A and B, are male nonheads of households in the Northeast region and equal in other characteristics, and A's occupation is lower on the census listing (has a higher code number) than B's, then chances are that A has a higher water related factor score than B.

In the North Central region and in the South also among males the S.P.'s occupation status coefficient is positive. In the West, although not significant, the coefficient is negative. This suggests that occupation has a different kind of influence in the West.

Hypothesizing and explaining with such a great amount of detail are very time consuming and usually lead to more doubts than convictions. Consequently, we will turn to a broader viewpoint. In the future it may be reasonable to expect that the level of participation in such activities as are closely dependent upon demographic variables, place of residence and health conditions will not change. Those activities which are more closely associated with occupation, education, and income might be more unstable.

If this be so, then the level of Y_3 scores, the active factor scores, will remain steady in the future. An aging of the population would depress the general level, but other influences from increasing presence of young children and diminishing numbers of rural backgrounds would be offsetting.

The level of passive pursuits may be expected to rise as the educational level does and also as the population ages.

The backwoods factor is only lightly dependent on any of the variables of the study, but it is at least as sensitive to the occupational, educational, and income variables as it is to the demographic type. Only among southern and western males does it appear that a rise in income may be followed by a rise in level of backwoods activity; elsewhere the backwoods scores may be expected to drop.

The water related activities scores need further study before even a tentative guess may be offered on their future course.